

May 1989

SPACING EXPERIMENTS AND CCT-PLOTS

- Statistical approach -

CENTRE TECHNIQUE FORESTIER TROPICAL
Département du Centre de Coopération Internationale
en Recherche Agronomique pour le Développement (CIRAD)
45 bis, avenue de la Belle-Gabrielle - 94736 NOGENT-sur-MARNE CEDEX
(FRANCE)

May 1989

SPACING EXPERIMENTS AND CCT-PLOTS

- Statistical approach -

CENTRE TECHNIQUE FORESTIER TROPICAL

Département du Centre de Coopération Internationale
en Recherche Agronomique pour le Développement (CIRAD)

45 bis, avenue de la Belle-Gabrielle - 94736 NOGENT-sur-MARNE CEDEX
(FRANCE)

Part 1 : THE NELDER

Part 2 : ANALYSIS OF A COMPACT DESIGN
FOR SPACING EXPERIMENTS

Part 3 : THE CCT-PLOTS

INTRODUCTION TO SPACING DESIGN (Statistical point of view)

When setting up a plantation, given a species, one of the first questions to answer to is : "What is the optimal distance that should separate two trees ?". This point is of great importance, because if the trees are set too close one from each other, the mortality rate that may affect them may be very high ; on the other hand, if the trees are too far from each other, their shape may not be good.

We shall describe and study two designs concerning spacing, respectively the Nelder design and the Marynen design.

We shall not discuss their qualities in the planter point of view, but only evoke the data processing.

THE NELDER

This design can be considered as a preliminary trial to further experiment that can bring interesting information. Its purpose is to study the behavior of trees set in different spacing conditions.

Principle : In the Nelder trial, the unit is the tree not the plot. The trees are located on concentric circles the rays of which are following a geometrical progression with α ratio.

$$r_n = \alpha r_{n-1} = \alpha r_0$$

The rays are equidistant with θ angle.

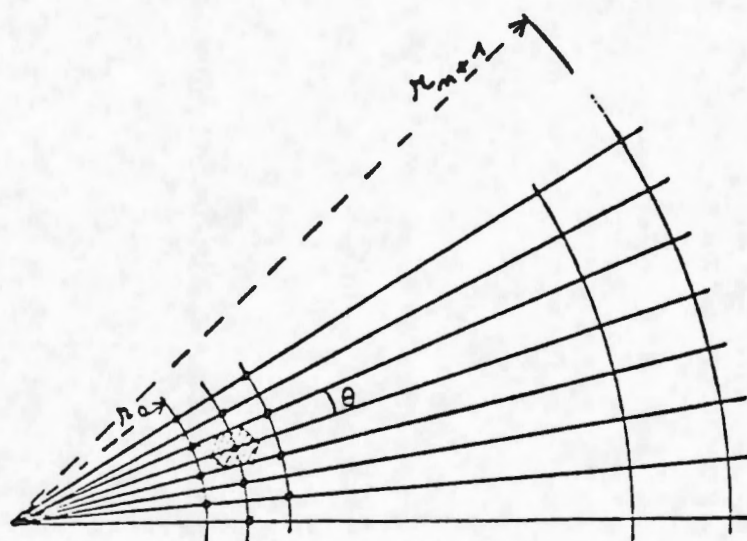
The trees are set staggered on one intersection out of two.

The tree situated at r_n distance of the center of the trial has a S_n "lebensaum" area at its disposal ; $S_n = \alpha^2 S_{n-1}$.

Determination of the characteristics of a Nelder trial :

Given n , the number of different spacings to be tested, S_1 the area corresponding to the smallest spacing, S_n the area corresponding to the greatest spacing, one must decide on the values of α , θ and r_0 with the following restriction on θ ; θ should take integer values only for practical reasons.

The trees situated on the circles the rays of which are r_0 and r_{n+1} as well as those located at the ends of the arcs of circles will be considered as border plants.



Value of α : the following formula is used : $\alpha^{2n-2} = \frac{S_n}{S_1}$.

In fact, the actual value of α will be close to the one calculated with the upper formula according θ to be integer.

From α and θ , we can calculate a parameter called μ which in fact is tabulated ; using μ , we can write $r_0 = \mu \sqrt{S_1}$.

When r_0 is known, it is actually not very difficult to compute r_1, r_2 and so on ; then the values of the different areas S_1, S_2, S_{n-1} are easy to determine.

Example : spacing trial on Eucalyptus :

Let us suppose there are $n=10$ different spacings to be tested, taking their values between $S_1 = 2.15 \text{ m}^2$ and $S_n = 16 \text{ m}^2$.

$$\text{Then } \frac{S_n}{S_1} = 7.111 \text{ so that } \text{Log}(\alpha) = \frac{\text{Log } 7.111}{2n-2} = 0.10898$$

$$\alpha = 1.1151$$

$$2\theta = \sqrt{\alpha - 2\cos \theta + \frac{1}{\alpha}}$$

$$\text{Then } 3^\circ \leq \theta \leq 4^\circ .$$

If we choose $\theta = 3^\circ$

$$\alpha = 1.09489$$

$$\mu = 9.36736$$

$$r_0 = \mu \sqrt{S_1} = 14.5 \text{ m}$$

if we choose $\theta = 4^\circ$

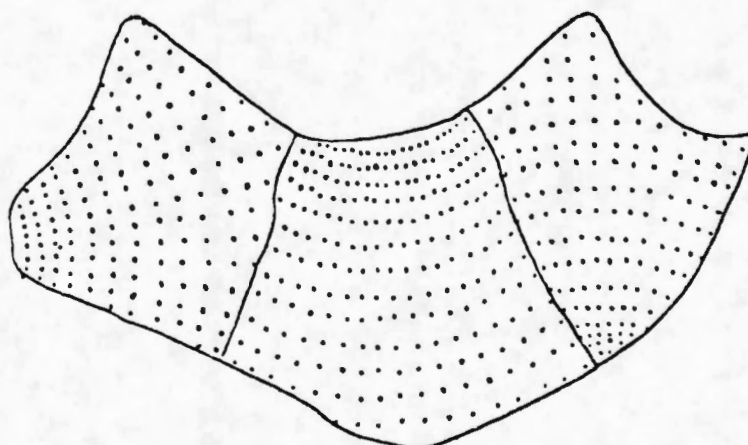
$$\alpha = 1.12842$$

$$\mu = 6.81647$$

$$r_0 = \mu \sqrt{S_1} = 10.22 \text{ m}$$

$r_n = \alpha r_{n-1}$	$S_n = \alpha^2 S_{n-1}$	$r_n = \alpha r_{n-1}$	$S_n = \alpha^2 S_{n-1}$
$r_1 = 15,38 \text{ m}$	$S_1 = 2,250 \text{ m}^2$	$r_1 = 11,53 \text{ m}$	$S_1 = 2,250 \text{ m}^2$
$r_2 = 16,84$	$S_2 = 2,697$	$r_2 = 13,02$	$S_2 = 2,865$
$r_3 = 18,44$	$S_3 = 3,233$	$r_3 = 14,69$	$S_3 = 3,648$
$r_4 = 20,19$	$S_4 = 3,876$	$r_4 = 16,57$	$S_4 = 4,645$
$r_5 = 22,11$	$S_5 = 4,647$	$r_5 = 18,70$	$S_5 = 5,915$
$r_6 = 24,21$	$S_6 = 5,570$	$r_6 = 21,10$	$S_6 = 7,532$
$r_7 = 26,50$	$S_7 = 6,678$	$r_7 = 23,81$	$S_7 = 9,590$
$r_8 = 29,02$	$S_8 = 8,005$	$r_8 = 26,87$	$S_8 = 12,212$
$r_9 = 31,77$	$S_9 = 9,596$	$r_9 = 30,32$	$S_9 = 15,549$
$r_{10} = 34,79$	$S_{10} = 11,504$	$r_{10} = 34,22$	$S_{10} = 19,800$
$r_{11} = 38,09$	$S_{11} = 13,791$	$r_{11} = 38,61$	$S_{11} = 25,211$
$r_{12} = 41,70$	$S_{12} = 16,532$		
$r_{13} = 45,66$	$S_{13} = 19,819$		

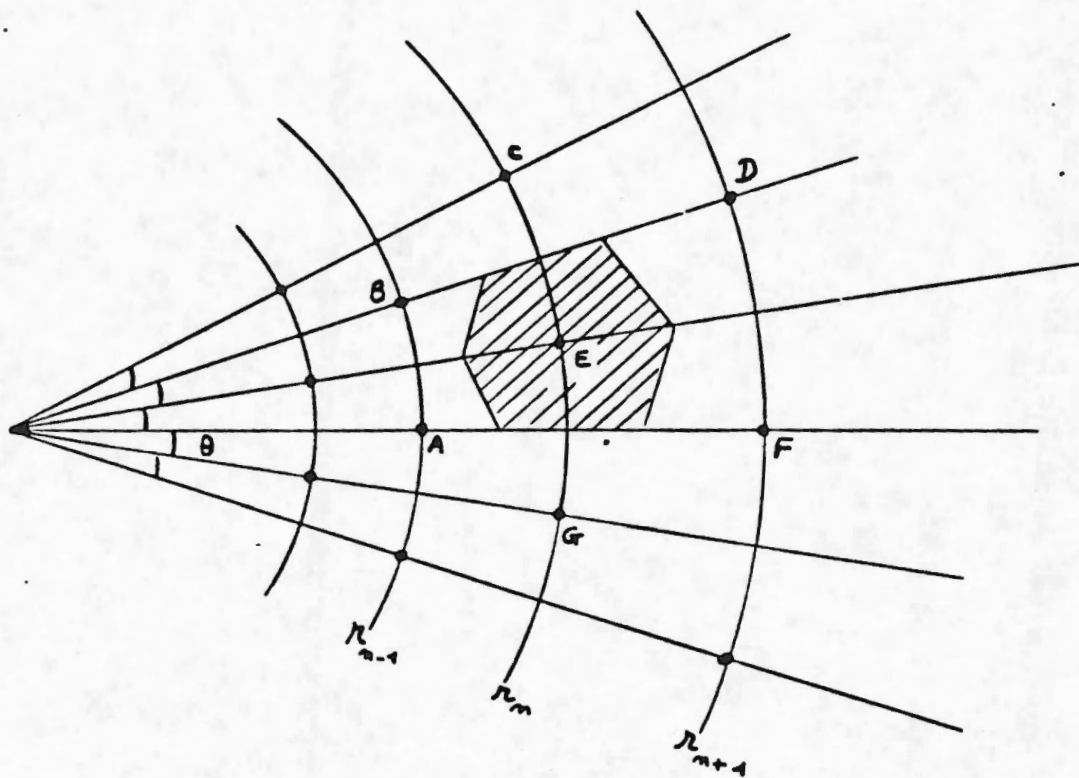
As in any other trial, soil heterogeneity should be controlled by using repetitions. Therefore the experiment ground should be divided in blocks while planting directions vary from one block to the other as illustrated below ; of course the proportion of border plants (non measured trees) is increased.



Brief indications about the analysis of the Nelder trial.

First, only the trees every neighbour of which is alive are kept (in this trial each tree except those forming the border lines has six neighbours : (see figure below)).

Tree "E" has six neighbours (namely "A", "B", "C", "D", "F", "G").



On the remaining population, the measured characteristic is plotted against the "lebenaum" area in order to locate the maximum.

DATA CORRESPONDING TO THE NELDER TRIAL
DESCRIBED PAGE BEFORE

Each column corresponds to a spacing distance, in fact a ray length.

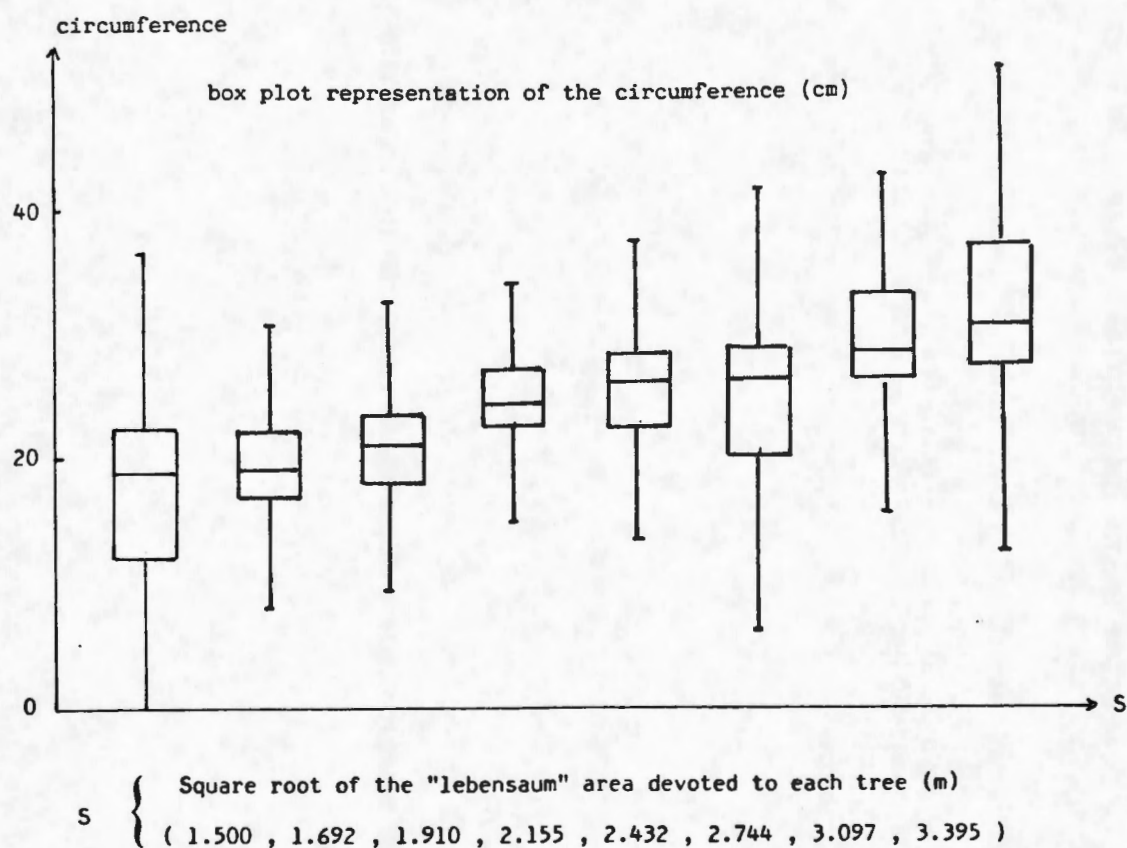
For instance, column 1 contains the values of the circumferences of the trees situated at a distance of 10.22 m from the center of the Nelder trial, and so on.

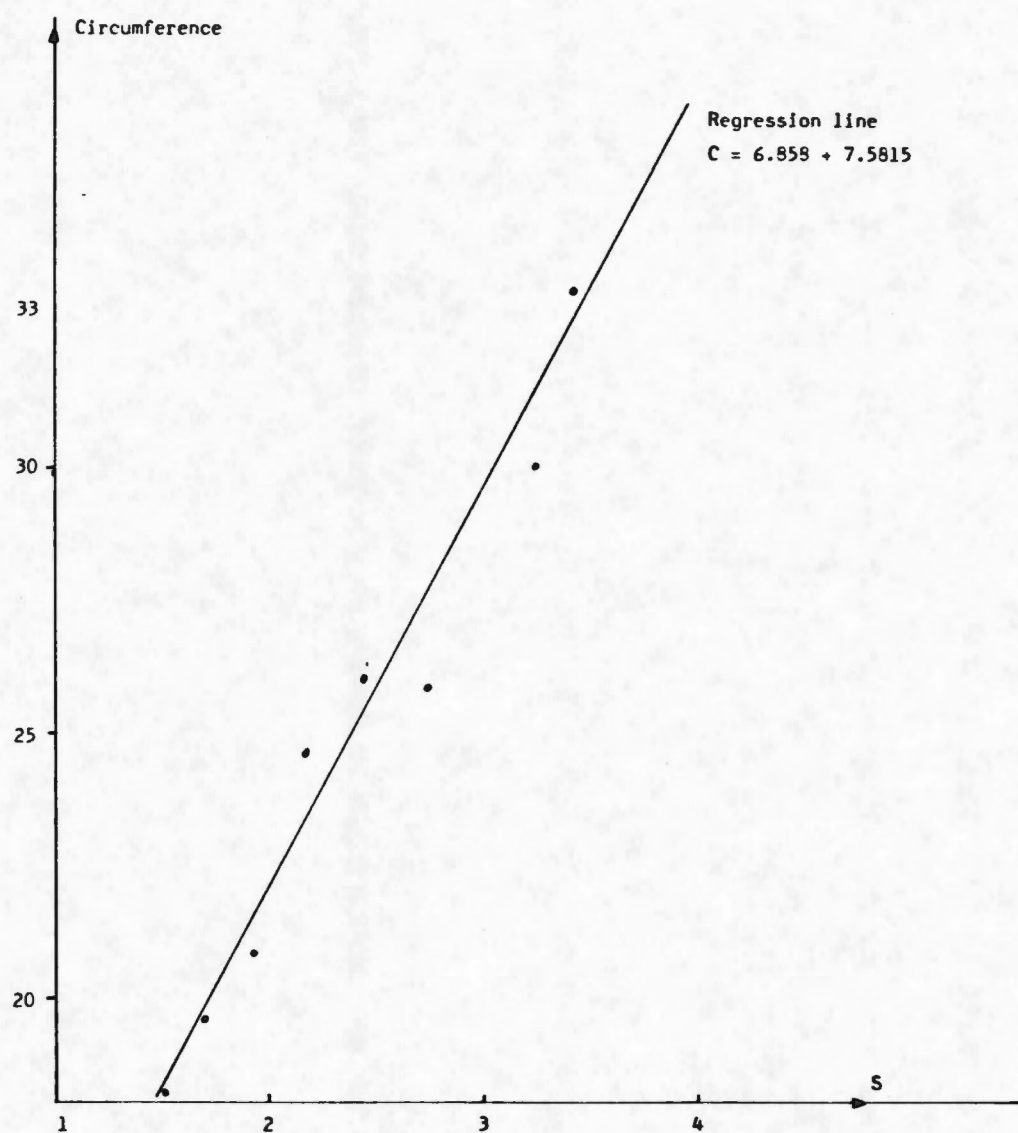
Circles :	1	2	3	4	5	6	7	8
	13	19	30	31	25	29	33	38
	13	24	20	20	31	15	27	32
	5	12	24	27	20	30	32	37
	20	23	22	24	28	22	31	36
	23	20	22	28	17	14	26	37
	19	17	28	19	29	29	30	41
	11	11	18	26	22	37	27	28
	9	29	17	10	27	34	32	40
	24	13	25	34	20	16	37	31
	20	20	11	32	30	41	38	42
	10	20	22	10	30	30	27	40
	42	19	18	25	26	17	35	32
	15	21	15	28	31	29	35	40
	22	23	20	19	29	21	36	26
	12	19	23	27	23	30	30	38
	11	19	26	24	27	25	27	31
	15	27	15	29	29	17	25	22
	25	12	21	32	17	24	27	32
	24	20	15	25	36	24	22	27
	22	17	27	23	28	27	23	25
	22	24	22	25	25	32	*	23
	20		18		26		*	

On the 7th column two trees have not been taken into account because of their bad behavior.

In order to treat the data with a box plot representation, we produce the following statistics concerning the circumference :

Column number	Number of trees	Mean	Standard error	Median	First quartile	Third quartile
1	22	18.045	1.6767	19.5	12	22
2	21	19.476	1.0411	20	17	23
3	22	20.863	1.0156	21.5	18	24
4	21	24.666	1.3875	25	23	28
5	22	26.181	1.0185	27	23	29
6	21	25.857	1.6172	27	21	30
7	20	30	1.0462	30	27	34
8	21	33.238	1.3695	32	28	38





ANALYSIS OF A COMPACT DESIGN FOR SPACING EXPERIMENTS

J.C. BERGONZINI
G. BURILLON

C.T.F.T./CIRAD
45 bis, avenue de la Belle-Gabrielle
94736 NOGENT-SUR-MARNE CEDEX - FRANCE

and

P. MOLIN
ENSBANA
Campus Universitaire - B.P. 138
21000 DIJON - FRANCE

June 1988

A. THE EXPERIMENTAL DESIGN

The purpose of this design is to study spacing effects on a quantitative variable concerning trees, which can be - depending on cases - the volume per hectare, the increase of basal area, the average girth, etc... (herein after referred to as yield).

The main advantage of this design is its compactness.

I - SETTING OF THE EXPERIMENT

The setting of the experiment is as follows :

given n spacings e_1, e_2, \dots, e_n , the treatment unit is formed by a plot of $R \times L$ trees arranged in L "lines" separated with e_i spacing and R "ranks" separated with e_j spacing. (See figure 1)

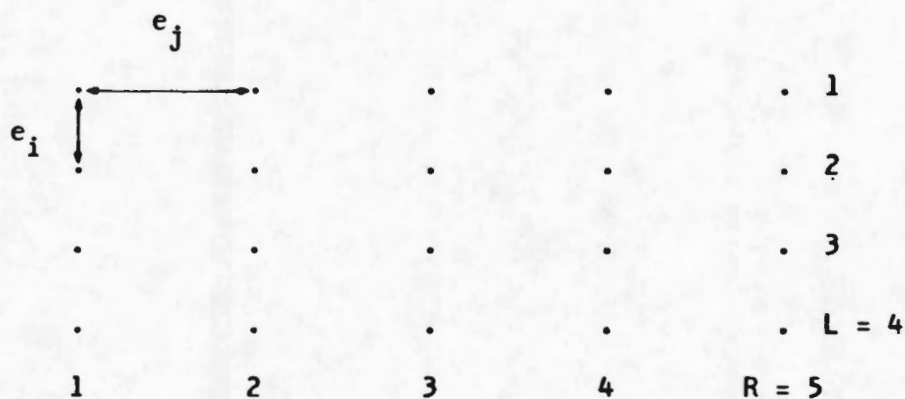


Figure 1

Each tree has an $e_i \times e_j$ "lebensaum" area at its disposal.

The n treatment units, obtained by carrying out all possible $e_i \times e_j$ crossings are then set in a rectangular block where all the plots with the same spacing according to either the lines or the ranks are arranged in "rows". (See figure 2)

.../...

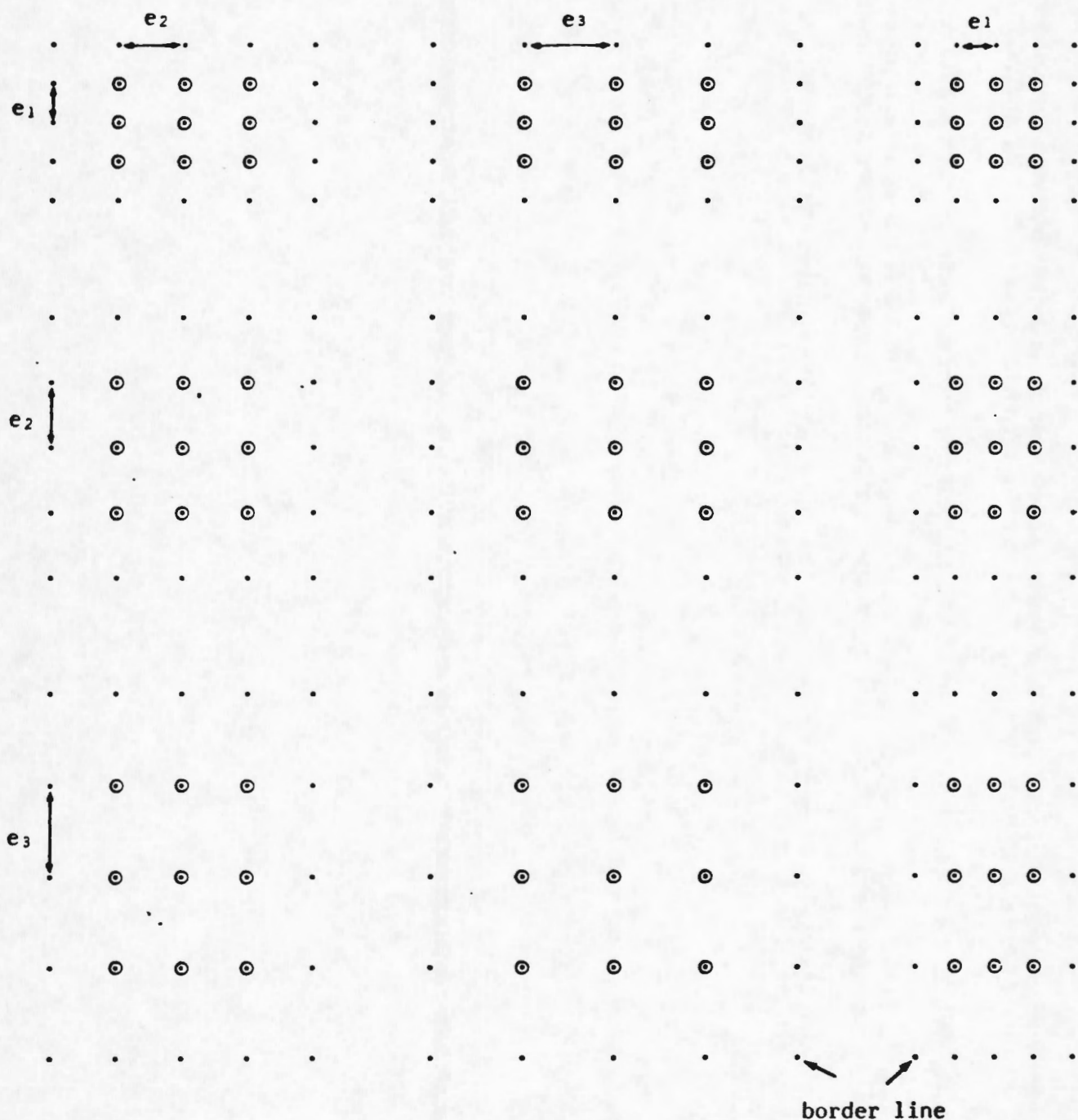


Figure 2 : Example of a design with three spacings and nine trees per plot.

Remark : The distribution of the spacings e_1 , e_2 , ..., e_n within the block is designated at random.

.../...

The global design is composed of b blocks ($b > 1$).
(example : see figure 3)

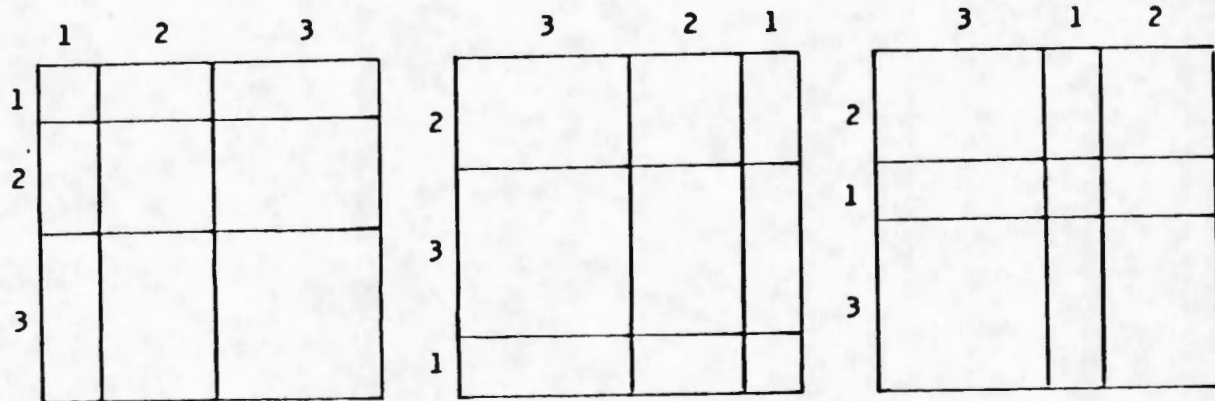


Figure 3 : Design with three blocks and three spacings.

- Remarks :
- a) Each plot comprises $R \times L$ measurable trees. One should not use $R \times L$ at less than 25 except in the case of particularly homogeneous clones.
 - b) The plots should be separated from the other by one or two border lines (or one or two ranks) ; (non-measured trees or rather trees whose measurements are not taken into account in the analysis).
 - c) It is better to set up square experimental units ($L = R$).
 - d) In order to respect the validity of the proposed analysis of variance model, it is necessary to avoid testing too large a spacing range (eg. : $e_1 = 0,5$ m and $e_n = 7$ m).
 - e) As is often the case, when working in the field of forest experimentation, each tree is measured, but the variable taken into account is an abstract résumé of the information (usually a mean value).
 - f) This design in the form of a criss cross is to test the effects of two factors (here : spacings on lines, and spacings on ranks) on the dependent variable (the yield).

.../...

NOTATIONS :

Let $1 \leq i \leq a$, $1 \leq j \leq a$, $1 \leq k \leq b$

- a) If x_{ijk} indicates an observation on the experimental unit where the spacings e_i and e_j are associated with the k^{th} block.

We shall denote :

$$\begin{aligned} x_{ij0} &= \sum_k x_{ijk} & x_{ij.} &= \frac{x_{ij0}}{b} \\ x_{i0k} &= \sum_j x_{ijk} & x_{i.k} &= \frac{x_{i0k}}{a} \\ x_{0jk} &= \sum_i x_{ijk} & x_{.jk} &= \frac{x_{0jk}}{a} \end{aligned}$$

using the same conventions,

$$\begin{aligned} x_{i00} &= \sum_{jk} x_{ijk} & x_{000} &= \sum_{ijk} x_{ijk} \\ x_{i..} &= \frac{x_{i00}}{ab} & x_{...} &= \frac{x_{000}}{a^2b} \end{aligned}$$

- b) Capital letters are used to represent random variables.

So X_{ijk} indicates the random variable whose value is x_{ijk} in the experiment.

- c) The expected value and the variance of a random variable X are respectively denoted $E(X)$, $Var(X)$, and the covariance between two random variables X and Y is represented by $Cov(X,Y)$.
- d) Greek letters (α, β, \dots) are generally used to denote unknown parameters whose estimates are written with a " $\hat{}$ "; for instance λ_{ij} is an unknown parameter estimated by $\hat{\lambda}_{ij}$.

.../...

B. THE MODELS

I - Introduction :

The model related to a criss cross design used to study the effects of two factors F and G on a quantitative variable X can be written as follows :

$$X_{ijk} = \mu + \alpha_i + \beta_j + \vartheta_{ij} + \delta_k + f_{ik} + g_{jk} + E_{ijk} \quad (1)$$

($1 \leq i \leq a_1$; $1 \leq j \leq a_2$ and $1 \leq k \leq b$) , where a_1 and a_2 are the respective numbers of modality of the two factors F and G and b is the number of blocks ; in the model that we are going to use here, the two factors F and G are identical, consequently :

$$a_1 = a_2 = a$$

$$\alpha_i = \beta_i = \text{effect of } e_i \text{ treatment}$$

$$\vartheta_{ij} = \text{interaction between } e_i \text{ treatment and } e_j \text{ treatment}$$

$$\delta_k = \text{effect of the } k^{\text{th}} \text{ block}$$

$$f_{ik} = \text{interaction between } e_i \text{ treatment and } k^{\text{th}} \text{ block}$$

$$g_{jk} = \text{interaction between } e_j \text{ treatment and } k^{\text{th}} \text{ block.}$$

As a matter of fact, since the two factors F and G are identical, it is not possible to differentiate one from the other within a block, so we have to transform the parameters f and g.

But, we can still choose either the "rank" or the "line" direction in each block ; then, if we denote :

$$\lambda_{ik} : \text{interaction between the } e_i \text{ spacing following the lines of the } k^{\text{th}} \text{ block}$$

$$\eta_{ik} : \text{interaction between the } e_i \text{ spacing following the ranks of the } k^{\text{th}} \text{ block and the } k^{\text{th}} \text{ block}$$

$$\lambda_{ik} = \begin{cases} f_{ik} \\ \text{or} \\ g_{ik} \end{cases} \quad \text{and} \quad \eta_{ik} = \begin{cases} f_{ik} & \text{if } \lambda_{ik} = g_{ik} \\ g_{ik} & \text{when } \lambda_{ik} = f_{ik} \end{cases}$$

.../...

we get :

$$\forall_k \quad \sum_i \lambda_{ik} = 0 \quad \text{and} \quad \sum_i \eta_{ik} = 0$$

while the extra conditions

$$\forall_i \quad \sum_k f_{ik} = 0 \quad \text{and} \quad \sum_k g_{ik} = 0$$

become

$$\forall_i \quad \sum_k (\lambda_{ik} + \eta_{ik}) = 0$$

II - Study of MODEL 1

- Model 1 can be written as follows :

$$X_{ijk} = \mu + \alpha_i + \alpha_j + \vartheta_{ij} + \delta_k + \lambda_{ik} + \eta_{jk} + E_{ijk}$$

$$(1 \leq i \leq a ; 1 \leq j \leq a ; 1 \leq k \leq b)$$

where the E_{ijk} are the independent and γ common variance random variables representing the zero mean residuals.

1. Meaning of the different parameters :

a) α_i represents the additive effect associated to the e_i spacing.

The α_i 's meet the conditions $\sum_{i=1}^a \alpha_i = 0$, so that there are
(a-1) independent α parameters.

b) δ_k represents the additive effect associated to the k^{th} block.

The δ_k 's meet the conditions $\sum_{k=1}^b \delta_k = 0$, so that there are
(b-1) independent δ parameters.

c) ϑ_{ij} is the interaction associated with the crossing of the e_i and e_j spacings.

.../...

The ϑ_{ij} 's meet the conditions :

$$\left. \begin{array}{l} \vartheta_{ij} = \vartheta_{ji} , \\ \forall_i , \sum_{j=1}^a \vartheta_{ij} = 0 \\ \forall_j , \sum_{i=1}^a \vartheta_{ij} = 0 \end{array} \right\} \text{ so that there are } \frac{a^2 - a}{2} \text{ independent } \vartheta \text{ parameters}$$

d) As explained above, λ_{ik} is the interaction between the e_i spacing as per the lines of the k^{th} block and the k^{th} block, and η_{ik} is the interaction between the e_i spacing as per the ranks of the k^{th} block and the k^{th} block, these parameters meet the following conditions :

$$\left. \begin{array}{l} \forall_k , \sum_{i=1}^a \lambda_{ik} = 0 \\ \forall_k , \sum_{j=1}^a \eta_{jk} = 0 \\ \forall_{ij} , \sum_{k=1}^b (\lambda_{ik} + \eta_{jk}) = 0 \end{array} \right\} \text{ so that there are } (2b-1)(a-1) \text{ independent } \lambda \text{ and } \eta \text{ parameters}$$

2. Estimation of the parameters :

The different parameters can be estimated by the classical method of the normalized equations associated with the linear model.

Let us set :

$$\Delta = \sum_{ijk} (x_{ijk} - \mu - \alpha_i - \alpha_j - \delta_k - \vartheta_{ij} - \lambda_{ik} - \eta_{jk})^2$$

2.1. Estimation of μ

$$\frac{\partial \Delta}{\partial \mu} = 0 \Rightarrow x_{000} - a^2 b \mu = 0 \Rightarrow \hat{\mu} = \bar{x} \dots$$

2.2. Estimation of α_1

$$\frac{\partial \Delta}{\partial \alpha_1} = 0 \Rightarrow \sum_{\substack{j,k \\ j \neq 1}} (x_{ijk} - \mu - \alpha_1 - \alpha_j - \delta_k - \vartheta_{ij} - \lambda_{ik} - \eta_{jk})$$

.../...

$$\begin{aligned}
 & + 2 \sum_k (x_{iik} - \mu - \alpha_i - \alpha_i - \delta_k - \vartheta_{ii} - \lambda_{ik} - \eta_{ik}) \\
 & + \sum_{\substack{j,k \\ j \neq i}} (x_{jik} - \mu - \alpha_j - \alpha_i - \delta_k - \vartheta_{ji} - \lambda_{jk} - \eta_{ik}) = 0 \\
 \Rightarrow \hat{\alpha}_i &= \frac{x_{i..} + x_{.i.}}{2} - \hat{\mu}
 \end{aligned}$$

2.3. Estimation of δ_k

$$\begin{aligned}
 \frac{\partial \Delta}{\partial \delta_k} = 0 &\Rightarrow \sum_{i,j} (x_{ijk} - \mu - \alpha_i - \alpha_j - \delta_k - \vartheta_{ij} - \lambda_{ik} - \eta_{jk}) = 0 \\
 &\Rightarrow \hat{\delta}_k = x_{..k} - \hat{\mu}
 \end{aligned}$$

2.4. Estimation of ϑ_{ij}

$$\text{When } i \neq j \quad \frac{\partial \Delta}{\partial \vartheta_{ij}} = 0 \Rightarrow \hat{\vartheta}_{ij} = \hat{\vartheta}_{ji} = \frac{x_{ij.} + x_{ji.}}{2} - \hat{\alpha}_i - \hat{\alpha}_j - \hat{\mu}$$

$$\text{When } j=i \quad \frac{\partial \Delta}{\partial \vartheta_{ii}} = 0 \Rightarrow \hat{\vartheta}_{ii} = x_{ii.} - \hat{\alpha}_i - \hat{\alpha}_i - \hat{\mu}$$

2.5. Estimation of λ_{ik} and η_{jk}

$$\frac{\partial \Delta}{\partial \lambda_{ik}} = 0 \Rightarrow \hat{\lambda}_{ik} = x_{i.k} - \hat{\alpha}_i - \hat{\delta}_k - \hat{\mu}$$

$$\frac{\partial \Delta}{\partial \eta_{jk}} = 0 \Rightarrow \hat{\eta}_{jk} = x_{.jk} - \hat{\alpha}_j - \hat{\delta}_k - \hat{\mu}$$

Remark : The condition $\sum_k (\hat{\lambda}_{ik} + \hat{\eta}_{jk}) = 0$ is verified for the two estimates here above given.

3. Splitting of $\sum_{ijk} x_{ijk}^2$ into sums of squares :

3.1. Contribution of the η_{jk} parameters

This contribution can be measured by the discrepancy between

$$\sum_{ijk} (x_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\alpha}_j - \hat{\delta}_k - \hat{\vartheta}_{ij} - \hat{\lambda}_{ik})^2 \text{ and } \Delta.$$

.../...

This difference equals $a \sum_{jk} \hat{\eta}_{jk}^2$

$$\text{so } \Delta = \sum_{ijk} (x_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\alpha}_j - \hat{\delta}_k - \hat{\vartheta}_{ij} - \hat{\lambda}_{ik})^2 - a \sum_{jk} \hat{\eta}_{jk}^2$$

3.2. Contribution of the λ_{ik} parameters

Using the same method (which applies to all kinds of parameters), we obtain the contribution of the $\hat{\lambda}_{ik}$ parameters equals

$$a \sum_{ik} \lambda_{ik}^2,$$

$$\text{so } \Delta = \sum_{ijk} (x_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\alpha}_j - \hat{\delta}_k - \hat{\vartheta}_{ij})^2 - a \sum_{ik} \hat{\lambda}_{ik}^2 - a \sum_{jk} \hat{\eta}_{jk}^2$$

3.3. Contribution of the ϑ_{ij} parameters

Δ can be written :

$$\Delta = \sum_{ijk} (x_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\alpha}_j - \hat{\delta}_k)^2 - b \sum_{ij} \hat{\vartheta}_{ij}^2 - a \sum_{ik} \hat{\lambda}_{ik}^2 - a \sum_{jk} \hat{\eta}_{jk}^2$$

so the contribution of the ϑ_{ij} parameters is

$$b \sum_{ij} \hat{\vartheta}_{ij}^2$$

3.4. Contribution of the δ_k parameters

Δ can be written :

$$\Delta = \sum_{ijk} (x_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\alpha}_j)^2 - a^2 \sum_k \hat{\delta}_k^2 - b \sum_{ij} \hat{\vartheta}_{ij}^2 - a \sum_{ik} \hat{\lambda}_{ik}^2 - a \sum_{jk} \hat{\eta}_{jk}^2$$

so the contribution of the δ_k parameters is

$$a^2 \sum_k \hat{\delta}_k^2$$

3.5. Contribution of the α_i parameters

Δ can be written :

$$\Delta = \sum_{ijk} (x_{ijk} - \hat{\mu})^2 - 2ab \sum_i \hat{\alpha}_i^2 - a^2 \sum_k \hat{\delta}_k^2 - b \sum_{ij} \hat{\vartheta}_{ij}^2 - a \sum_{ik} \hat{\lambda}_{ik}^2 - a \sum_{jk} \hat{\eta}_{jk}^2$$

so the contribution of the α_i parameters is

$$2ab\sum_i \hat{\alpha}_i^2$$

3.6. Contribution of the μ parameter =

(global splitting of $\sum_{ijk} x_{ijk}^2$ into sums of squares).

$$\sum_{ijk} x_{ijk}^2 = a^2 b \hat{\mu}^2 + 2ab \sum_i \hat{\alpha}_i^2 + a^2 \sum_k \hat{\delta}_k^2 + b \sum_{ij} \hat{\vartheta}_{ij}^2 + a \sum_{ik} \hat{\lambda}_{ik}^2 + a \sum_{jk} \hat{\eta}_{jk}^2 + \Delta$$

we can see that the contribution of the μ parameter is

$$a^2 b \hat{\mu}^2$$

3.7. From another point of view :

In the $\mathbb{R}^{a^2 b}$ vector space the observation vector is written

$$\vec{x} = \sum_{ijk} x_{ijk} \vec{e}_{ijk} \quad \text{where } \vec{e}_{ijk} \text{ is a canonical basis.}$$

One can consider the vector subspaces associated with the different parameters

- E_1 is the subspace generated by $\sum_{ijk} \mu \vec{e}_{ijk}$ ($\dim E_1 = 1$)
- E_2 is the subspace generated by $\sum_{ijk} \alpha_i \vec{e}_{ijk}$ ($\dim E_2 = a-1$)
- E_3 is the subspace generated by $\sum_{ijk} \alpha_j \vec{e}_{ijk}$ ($\dim E_3 = a-1$)
- E_4 is the subspace generated by $\sum_{ijk} \delta_k \vec{e}_{ijk}$ ($\dim E_4 = b-1$)
- E_5 is the subspace generated by $\sum_{ijk} \vartheta_{ij} \vec{e}_{ijk}$ ($\dim E_5 = a \frac{(a-1)}{2}$)
- E_6 is the subspace generated by $\sum_{ijk} \lambda_{ik} \vec{e}_{ijk}$
($\dim E_6 = (2b-1)(a-1)$)
- E_7 is the subspace generated by $\sum_{ijk} \eta_{jk} \vec{e}_{ijk}$
($\dim E_7 = (2b-1)(a-1)$)

In fact, because of the particular behaviour of the two factors, we are interested in $E_{(2,3)}$ subspace generated by

$$\sum_{ijk} (\alpha_i + \alpha_j) \vec{e}_{ijk}$$

.../...

and $E_{(6,7)}$ subspace generated by

$$\sum_{ijk} (\lambda_{ik} + \eta_{jk}) \vec{e}_{ijk}$$

Then, if \mathbb{R}^{a^2b} is given a metric associated with the identity matrix I it can be proved that :

- E_2 and E_3 are orthogonal as well as E_6 and E_7 ,
- $\dim E_{(2,3)} = \dim E_2 = a-1$,
- $\dim E_{(6,7)} = \dim E_6 = (2b-1)(a-1)$.

Finally, it can be proved that the subspaces E_1 , $E_{(2,3)}$, E_5 and $E_{(6,7)}$ are orthogonal and this bears out the splitting of

$$\sum_{ijk} x_{ijk}^2$$

III - Back to the model - Distribution of the estimates and of the sums of squares - MODEL 2

Since the spacings are distributed at random among the lines and the ranks, the λ_{ik} and η_{jk} can be considered as being the values taken by random variables, respectively P_{ik} and Q_{jk} . The same remark applies to the block effect which can be regarded as random or fixed.

So, the model we shall consider in the following can be written (block effect is fixed) :

Model 2

$$X_{ijk} = \mu + \alpha_i + \alpha_j + \vartheta_{ij} + \delta_k + P_{ik} + Q_{jk} + E_{ijk}$$

with the same conditions still applying to α , ϑ and δ , plus extra conditions to P , Q and E :

$$\forall_{i,k} \quad E(P_{ik}) = 0 \quad \text{Var} (P_{ik}) = \gamma_I$$

$$\forall_{i,k} \quad E(Q_{jk}) = 0 \quad \text{Var} (Q_{jk}) = \gamma_I$$

$$\forall_{i,j,k} \quad E(E_{ijk}) = 0 \quad \text{Var} (E_{ijk}) = \gamma$$

and the variables P_{ik} , Q_{jk} and E_{ijk} being independent.

.../...

1. Distribution of $X_{...}$ as an estimate of μ

$$X_{...} = \mu + P_{..} + Q_{..} + E_{...}$$

$$E(X_{...}) = \mu$$

$$\text{Var}(X_{...}) = \frac{1}{a^2 b} (2a\gamma_I + \gamma)$$

So, $\frac{a^2 b X_{...}^2}{2a\gamma_I + \gamma}$ has a χ'^2 distribution with one degree of freedom and a non centrality parameter $\lambda = \frac{a^2 b \mu^2}{2a\gamma_I + \gamma}$

2. Distribution of A_i as an estimate of α_i

$$A_i = \alpha_i + \frac{P_{i.} - P_{..}}{2} + \frac{Q_{i.} - Q_{..}}{2} + \frac{E_{i..} - E_{...}}{2} + \frac{E_{.i.} - E_{...}}{2}$$

$$E(A_i) = \alpha_i ,$$

$$\text{Var}(A_i) = \frac{1}{a^2 b} (a-1)(a\gamma_I + \gamma) ,$$

$$\text{Cov}(A_i, A_{i'}) = - \frac{1}{2a^2 b} (a\gamma_I + \gamma)$$

Under the null hypothesis H_0 ($\forall_i, \alpha_i = 0$) the statistic

$$\frac{2ab \sum_{i=1}^a A_i^2}{a\gamma_I + \gamma}$$

usually used to test the presence of treatment effects has a χ^2 distribution with $(a-1)$ degrees of freedom.

3. Distribution of B_k as an estimate of δ_k

$$B_k = \delta_k + (P_{.k} - P_{..}) + (Q_{.k} - Q_{..}) + (E_{...k} - E_{...})$$

$$E(B_k) = \delta_k$$

$$\text{Var}(B_k) = \frac{b-1}{a^2 b} (2a\gamma_I + \gamma)$$

$$\text{Cov}(B_k, B_{k'}) = - \frac{1}{a^2 b} (2a\gamma_I + \gamma)$$

.../...

The classical estimate of the block effect is

$$\frac{a^2 \sum_{k=1}^b B_k^2}{2a \gamma_I + \gamma}$$

and under the null hypothesis ($H_0 : \forall_k ; \delta_k = 0$), it has a χ^2 distribution with $(b-1)$ degrees of freedom.

4. Distribution of θ_{ij} as an estimate of ϑ_{ij}

When $i \neq j$,

$$\theta_{ij} = \vartheta_{ij} + \frac{1}{2} (E_{ij.} - E_{i..} - E_{.j.} + E_{...}) + \frac{1}{2} (E_{j.i.} - E_{j..} - E_{.i.} + E_{...})$$

When $i = j$,

$$\theta_{ii} = \vartheta_{ii} + (E_{ii.} - E_{i..} - E_{.i.} + E_{...})$$

$$\forall_{i,j}, E(\theta_{ij}) = \vartheta_{ij}$$

When $i \neq j$,

$$\text{Var}(\theta_{ij}) = \frac{\gamma}{2a^2b} [(a-1)^2 + 1]$$

When $i = j$,

$$\text{Var}(\theta_{ii}) = \frac{\gamma}{b} \frac{(a-1)^2}{a^2}$$

In order to express the distribution of

$$\frac{b \sum_{ij} \theta_{ij}^2}{\gamma},$$

we first need to know the covariances of the θ_{ij} .

When $i \neq i' ; j \neq j' ; i \neq j'$ and $i' \neq j$

$$\text{Cov}(\theta_{ij} ; \theta_{i'j'}) = \frac{\gamma}{a^2b},$$

When $i = i' ; j \neq j' ; i \neq j$ and $i \neq j'$

$$\text{Cov}(\theta_{ij} ; \theta_{ij'}) = \frac{1}{2} \frac{\gamma}{a^2b} (a+2)$$

(same result when $j' = i$ and $i' = j'$, for symmetrical reasons).

.../...

When $p \neq i$; $m \neq i$; $p \neq m$

$$\text{Cov} (\theta_{ii} ; \theta_{pm}) = \frac{\gamma}{a^2 b} ,$$

When $p \neq i$,

$$\text{Cov} (\theta_{ii} ; \theta_{pp}) = \frac{\gamma}{a^2 b} ,$$

When $p \neq i$,

$$\text{Cov} (\theta_{ii} ; \theta_{ip}) = - \frac{(a-1) \gamma}{a^2 b} ,$$

Then, under the null hypothesis H_0 , $(\forall_{ij} \theta_{ij} = 0)$,

$$E (b \sum_{ij} \theta_{ij}^2) = \frac{\gamma a(a-1)}{2}$$

and $\frac{b \sum_{ij} \theta_{ij}^2}{\gamma}$ has a χ^2 distribution with $\frac{a(a-1)}{2}$ degrees of freedom.

5. Distribution of the Λ_{ik} and H_{jk} as estimates of λ_{ik} and η_{jk} (parameters of model 1)

$$\Lambda_{ik} = P_{ik} - P_{.k} - \frac{P_{i.} - P_{..}}{2} - \frac{Q_{i.} - Q_{..}}{2} + E_{i.k} - E_{..k} - \frac{E_{i..} - E_{...}}{2} - \frac{E_{.i.} - E_{...}}{2}$$

$$E (\Lambda_{ik}) = 0$$

$$\text{Var} (\Lambda_{ik}) = \frac{(a-1)(2b-1)}{2a^2 b} (a\gamma_I + \gamma)$$

The same results apply to H_{jk} .

But, what is of greater interest here is the distribution of

$$a \sum_{ik} \Lambda_{ik}^2 + a \sum_{jk} H_{jk}^2$$

which is the sum of squares contributing to the splitting of

$$\sum_{ijk} x_{ijk}^2$$

$$a \sum_{ik} \Lambda_{ik}^2 + a \sum_{jk} H_{jk}^2 = a \sum_{ik} (X_{i.k} - X_{i..} - X_{..k} + X_{...})^2 + a \sum_{jk} (X_{.jk} - X_{.j.} - X_{..k} + X_{...})^2$$

$$+ 2ab \sum_i \left(\frac{X_{i..} - X_{.i.}}{2} \right)^2$$

.../...

Under the null hypothesis (absence of interaction)

both $a \sum_{ik} (X_{i.k} - X_{i..} - X_{..k} + X_{...})^2$ and $a \sum_{jk} (X_{.jk} - X_{.j.} - X_{..k} + X_{...})^2$

have distributions proportional to χ^2 with $(a-1)(b-1)$ degrees of freedom while

$$2ab \sum_i \left(\frac{X_{i..} - X_{..i.}}{2} \right)^2$$

has a distribution proportional to a χ^2 with $a-1$ degrees of freedom. These three sums of squares are independent, so

$$\frac{a \sum_{ik} \Lambda_{ik}^2 + a \sum_{jk} H_{jk}^2}{a \gamma_I + \gamma}$$

has a χ^2 distribution with $(a-1)(2b-1)$ degrees of freedom.

6. Distribution of the residuals R_{ijk}

$$\begin{aligned} R_{ijk} &= X_{ijk} - X_{...} - A_i - A_j - B_k - \theta_{ij} - \Lambda_{ik} - H_{jk} \\ &= E_{ijk} - \frac{1}{2}(E_{ij.} - E_{i..} - E_{.j.} + E_{...}) - \frac{1}{2}(E_{ij.} - E_{j..} - E_{.i.} + E_{...}) - E_{i.k} - E_{.jk} + E_{...} \end{aligned}$$

R_{ijk} is a function of E_{ijk} only, then similar to the one obtained in the case of model 1.

Then

$$\frac{\sum_{ijk} R_{ijk}}{\gamma}$$

follows a χ^2 distribution with $(b(a-1)^2 - \frac{(a-1)a}{2})$ degrees of freedom.

.../...

IV - Analysis of variance

These results lead to the following table of analysis of variance :

Sources of variations	Sums of squares	Degrees of freedom	Meaned squares	Expected values/H
Spacing	$2ab \sum_i A_i^2$	$(a-1)$	$\frac{2ab \sum_i A_i^2}{a-1}$ (1)	$a \gamma_I + \gamma$
Block	$a^2 \sum_k B_k^2$	$(b-1)$	$\frac{a^2 \sum_k B_k^2}{b-1}$ (*)	$2a\gamma_I + \gamma$
Interaction between spacings	$b \sum_{ij} \theta_{ij}^2$	$\frac{a(a-1)}{2}$	$\frac{2b \sum_{ij} \theta_{ij}^2}{a(a-1)}$ (2)	γ
Block-spacing interaction	$a \sum_{ik} \Lambda_{ik}^2 + a \sum_{jk} H_{jk}^2$	$(a-1)(2b-1)$	$\frac{a \sum_{ijk} (\Lambda_{ik} + H_{jk})^2}{(a-1)(2b-1)}$ (3)	$a \gamma_I + \gamma$
Residual	Δ	$b(a-1)^2 - \frac{a(a-1)}{2}$	$\frac{\Delta}{b(a-1)^2 - \frac{a(a-1)}{2}}$ (4)	γ

* the block effect cannot be tested.

- In order to test the spacing effect the ratio (1)/(3), which under the null hypothesis (no spacing effect) has a FISHER - SNEDECOR distribution with $(a-1)$ and $(a-1)(2b-1)$ degrees of freedom, should be considered.
- In order to test the presence of the interaction between the spacings the ratio (2)/(4), which under the null hypothesis (lack of interaction) follows a FISHER - SNEDECOR distribution with $(\frac{a(a-1)}{2})$ and $(b(a-1)^2 - \frac{a(a-1)}{2})$ degrees of freedom, should be considered.

N.B. : The presence of interaction should be tested first, then in the case of lack of interaction the treatment effect should be tested.

.../...

V - Comparisons by pairs

1. Comparison of the A_i 's

In the case no θ_{ij} interaction is detected but a treatment effect becomes apparent (\exists_i such as $\alpha_i \neq 0$) then the different spacings can be compared by pairs, for instance by means of a Bonferroni test.

$A_i - A_{i'}$ is normally distributed with true mean $\alpha_i - \alpha_{i'}$ and variance $\frac{1}{ab} (a-2) (a\gamma_I + \gamma)$.

To compare two spacings is to test the nullity of the expected mean of $A_i - A_{i'}$. It has been assessed (see Analysis of variance table) that $a\gamma_I + \gamma$ can be estimated by :

$$V = \frac{a \sum (\Lambda_{ik} + H_{jk})^2}{(a-1)(2b-1)}$$

and that A_i and V are independent variables.

Then, under the null hypothesis, ($H_0 ; \alpha_i = \alpha_{i'}$) the statistic

$$T_{ii'} = \frac{A_i - A_{i'}}{\sqrt{\frac{a-2}{b} \frac{1}{(a-1)(2b-1)} \sum_{ijk} (\Lambda_{ik} + H_{jk})^2}}$$

has a Student distribution with $(a-1)(2b-1)$ degrees of freedom.

Remark : If one wishes to compare all pairs of spacings, one should choose, for each comparison, a confidence probability equal to :

$$\frac{\alpha}{C_a^2} = \frac{2\alpha}{a(a-1)} \quad (\text{where } \alpha \text{ is the confidence probability used in the analysis of variance}).$$

2. Comparison of the θ_{ij} 's

When there is an interaction effect, one may want to compare the different spacing crossings $e_i \times e_j$. The techniques given in V - 1. do apply, but interpreting the results is far more difficult.

.../...

Example :

The following design is related to a spacing experiment which was carried out at the Station of LOANDJILI (POINTE-NOIRE - CONGO).

Description :

a) Six spacings are to be tested :

$e_1 = 3,5 \text{ m}$ $e_2 = 4 \text{ m}$ $e_3 = 4,5 \text{ m}$ $e_4 = 5 \text{ m}$
 $e_5 = 5,5 \text{ m}$ $e_6 = 6 \text{ m}$.

b) There are nine trees per plot (which is rather few), and three blocks.

c) The studied species is the Eucalyptus PF1 - Clone 1.45 .

d) The dependent variable is the basal area per hectare at the age of fifty six months.

.../...

Tables of observations (The variable is the basal area/ha measured in mm²).

BLOCK 1

<div> <div>Lines</div> <div>Ranks</div> </div>		Spacings					
		3,5	4	4,5	5	5,5	6
Spacings	3,5	134,8	162,7	182,0	185,1	204,8	212,0
	4	152,7	185,9	199,6	216,25	231,3	243,1
	4,5	177,0	178,4	205,7	208,8	246,9	229,5
	5	175,8	195,3	224,0	232,7	242,0	255,9
	5,5	179,1	192,3	248,9	263,0	266,3	276,2
	6	206,6	185,3	244,0	271,3	287,9	297,5

BLOCK 2

<div> <div>Lines</div> <div>Ranks</div> </div>		Spacings					
		3,5	4	4,5	5	5,5	6
Spacings	3,5	145,1	164,7	161,8	202,2	189,6	206,6
	4	157,4	179,7	207,5	227,9	225,9	232,5
	4,5	174,4	189,1	225,1	248,1	234,6	263,0
	5	179,8	194,4	220,5	226,0	241,2	256,6
	5,5	185,75	231,1	233,5	281,6	280,7	275,9
	6	199,2	230,5	243,6	281,7	284,7	295,2

BLOCK 3

<div> <div>Lines</div> <div>Ranks</div> </div>		Spacings					
		3,5	4	4,5	5	5,5	6
Spacings	3,5	146,5	170,8	193,3	189,9	209,9	216,7
	4	168,1	161,4	192,5	210,4	212,7	208,1
	4,5	182,5	196,3	212,4	228,9	242,9	240,0
	5	171,5	195,1	228,8	224,6	249,9	263,3
	5,5	211,0	227,9	238,6	266,0	243,0	281,3
	6	214,3	235,6	248,7	266,3	289,6	291,1

Applying the calculations used earlier (computation of the different sums of squares such as $2ab \sum_{i=1}^a A_i^2$ etc...) leads to the following table of analysis of variances :

Sources of variation	Sums of squares	Degrees of freedom	Meaned squares	F
Spacing	142 421	5	26 484	112 *
Interaction between spacings	2 204,74	15	146,98	1,56
Block spacing interaction	6 352,95	25	254,11	
Residual	5 563,08	60	93,05	

* One concludes the presence of a spacing effect.

Comparison of the spacings (numerical example)

In our case :

Number of different spacings = $a = 6$; number of blocks = $b = 3$.

The confidence probability associated with the comparison of pairs of treatments = $\frac{2 \times 0,05}{6 \times 5} = 0,003$ (assuming that the probability confidence for the analysis of variance $\alpha = 5\%$)

and

$$\sqrt{\frac{a-2}{ab} \frac{a}{(a-1)(2b-1)} \sum (\hat{\lambda}_{ik} + \hat{\eta}_{jk})^2} = \sqrt{\frac{4}{6 \times 3} 254,11} = 7,51$$

then, the differences $\alpha_i - \alpha_1$, should be compared to $7,51 \times 3,1 = 23,3$

then, if we regard e_4 and e_6 they have to be considered as different

since $\alpha_6 = 33,75$ $\alpha_4 = 8,48$ ($\alpha_6 - \alpha_4 = 25,27 > 23,3$) .

.../...

APPENDIX :

Below are some results that are useful to determine the distributions of the different estimates of model 2.

A) a) Given Z_1, \dots, Z_n , n independent identically distributed variables ($\text{Var } Z_i = \sigma^2$; \forall_i)

$$\text{then } \text{Var} (Z_i - Z_1) = \frac{n-1}{n} \sigma^2$$

b) Given Z_{ij} ($1 \leq i \leq n$; $1 \leq j \leq m$), nm independent identically distributed variables

$$\text{then } \text{Cov} (Z_{i.} - Z_{..} ; E_{ij} - Z_{..}) = 0$$

and when $i \neq i'$

$$\text{Cov} (Z_{i.} - Z_{..} ; Z_{i'.} - Z_{..}) = \frac{-\sigma^2}{nm}$$

B) If X is a random normally distributed vector ($X \sim N(\mu, V)$),

assuming that :

$$1) VAVAV = VAV$$

$$2) \mu'AV = \mu'AVAV$$

$$3) \mu'AV\mu = \mu'AVAV\mu ,$$

it can be proved that $X'AX$ follows a χ'^2 distribution whose number of degrees of freedom is $\text{tr}(AV)$ and whose non-centrality parameter equals $\frac{1}{2} \mu'AV\mu$.

C) If Z_{ij} ($1 \leq i \leq n$; $1 \leq j \leq m$) are nm independent identically distributed variables ($\forall_{ij} Z_{ij} = \sigma^2$)

then :

$$a) \text{Var} (Z_{ij} - Z_{i.} - Z_{.j} + Z_{..}) = (1 - \frac{1}{n})(1 - \frac{1}{m}) \sigma^2$$

b) When $i \neq i'$; $j \neq j'$,

$$\text{Cov} (Z_{ij} - Z_{i.} - Z_{.j} + Z_{..} ; Z_{i'j'} - Z_{i'.} - Z_{.j'} + Z_{..}) = \frac{\sigma^2}{nm}$$

.../...

REFERENCES :

BERGONZINI, J.C. (1984) - Proposition pour le traitement d'un essai écartement - CTFT/CIRAD, France.

COCHRAN, W.G. and COX, G.M. (1957) - Experimental designs - 2nd ed. - WILEY, New York.

CHUAN-SHENG LIN and MORSE, P.M. - A compact design for spacing experiments (BIOMETRICS 31, 661-671 - Sept. 1975)

NELDER, J.A. - New kinds of systematic designs for spacing experiments (BIOMETRICS 18, 283-307 - 1962)

SEARLE, S.R. (1971) - Linear models -
WILEY, New York.

°
° °

c) When $j \neq j'$,

$$\text{Cov} (Z_{ij} - Z_{i.} - Z_{.j} + Z_{..} ; Z_{ij'} - Z_{i.} - Z_{.j'} + Z_{..}) = (1 - \frac{1}{n}) \frac{\sigma^2}{m}$$

d) When $i \neq i'$,

$$\text{Cov} (Z_{ij} - Z_{i.} - Z_{.j} + Z_{..} ; Z_{i'j} - Z_{i'.} - Z_{.j} + Z_{..}) = (1 - \frac{1}{m}) \frac{\sigma^2}{n} .$$

° ° °

SOME STATISTICAL CONSIDERATIONS ABOUT A CLEARING DESIGN :
THE CCT-PLOTS

In plantations, a few years after having been planted, the trees have to face competition ; in order to let them grow competition free, the planter may use clearing as a mean to reaffect new spacings to the trees to control the growth of the trees and therefore the yield of the plantation.

We shall study a design called the CCT plots as a mean to point out the emergence of differences between the behaviors of two populations set at different densities.

About correlated curves trend plots : (CCT plots)

CCT plots is a forestry experimental design which was introduced in the thirties by O'CONNOR in South Africa. It is used in plantations as a mean to establish how many trees can grow together with competition free at successive degrees of their development which is measured in age or height. The aim of the CCT plots is to point out the effects of clearing.

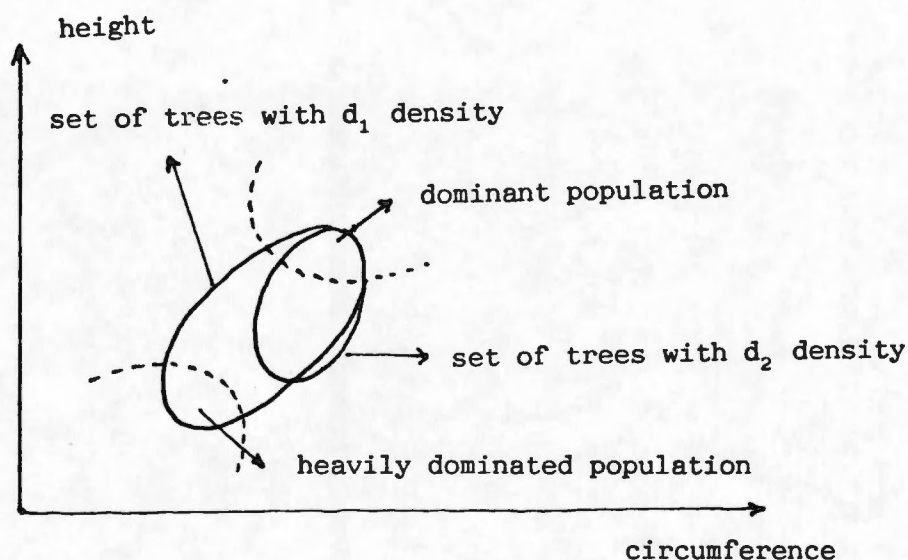
Mainly, the CCT plots are based on the setting of a competition free growing curve to be compared to individual growing in plots first submitted to competition then brought to densities of those which are free growing. They are a mean to estimate the size of the mean tree of the principal population on which some treatment is applied. But first, we have to describe the effects of competition and of clearing.

I - What are the effects of competition and the consequences of a clearing :

Given a population, two factors can modify its behavior : competition and clearing.

a) Competition :

Given two plots in the same conditions (soil fertility for instance), except for their densities (respectively d_1 and d_2 with $d_1 > d_2$), the competition effect can for a start be represented with the following figure.



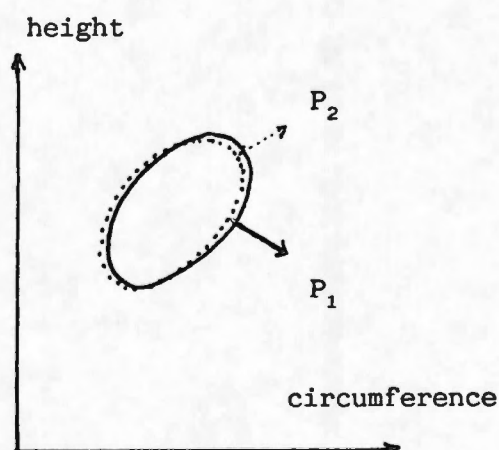
The increase of competition can be described with the following results :

- * no or very little effect on dominant population,
- * appearance of a heavily dominated population,
- * "regression" of the mean population.

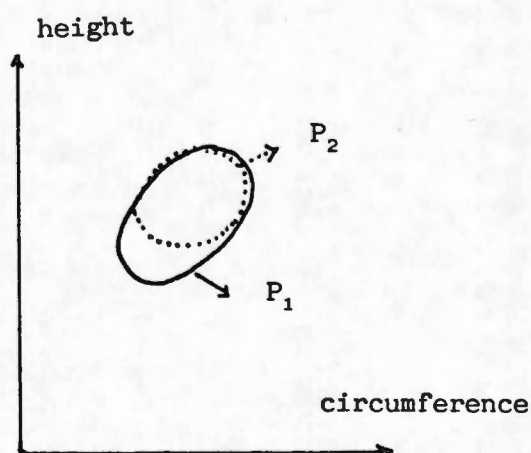
b) Clearing :

Given two identical populations, say P_1 and P_2 , selective clearing on P_2 has an effect which can be described with the following figures :

before clearing



after clearing applied on P_2



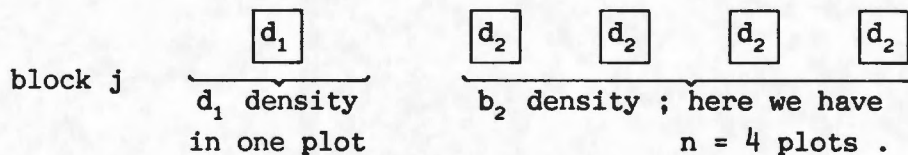
Consequences : When one wishes to compare two populations P_1 and P_2 . Let us suppose that a few years ago P_2 has been cleared more heavily than P_1 . The situation will be similar to the one described with the first figure (concerning competition), except that its characteristics will be more marked because of the clearing.

So, in order to compare the two populations concerning the effects of competition, it should be clever to consider the mean populations.

In an actual case : when one clears on the P_2 population, one should make an imaginary clearing on the P_1 population with the same intensity (one marks the trees which would have been eliminated). Let us denote P'_1 this fictitious population resulting of P_1 . Before performing the comparison between P'_1 and P_2 , the dominant populations are put apart : (say the 100 biggest trees of an hectare). Then, the remaining mean populations are compared.

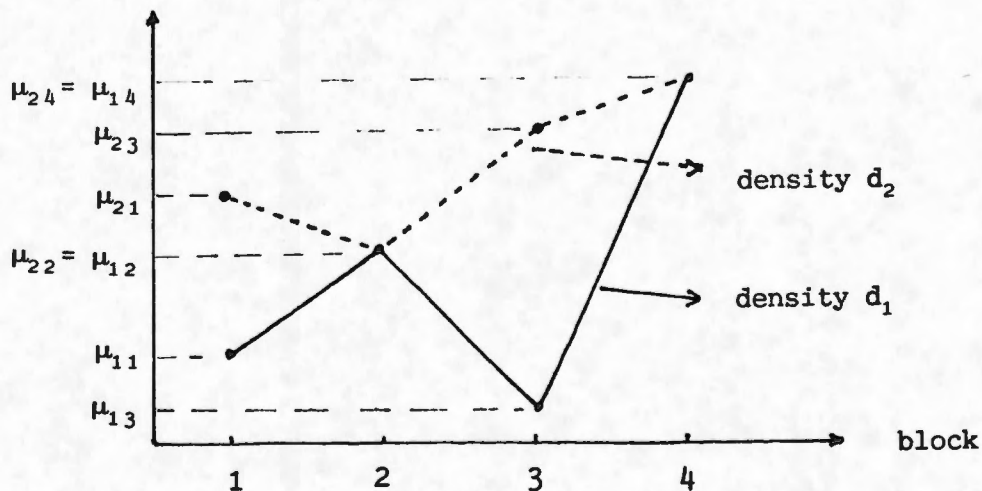
II - The analysis of variance of the CCT-plots : when should the clearing be decided ?

Usually the design is composed of b blocs and there are two densities to be compared. One of them is present once in a block, the other is repeated n times in a block.



If the blocks are of different richness of soil, the following problem arises : the competition might have appeared in some blocks and not in the others.

Mean result of density i on the block $j = \mu_{ij}$



Example : Four blocks and two densities ; here competition has broken out in blocks number one and number three ; the consequences of fertility is the possible presence of an interaction "block x density".

Model : Let us denote X , the measured variable (circumference, height, increase,...), X_{ijk} the random variable associated to the i density, the j^{th} block and the k^{th} repetition.

$$1 \leq i \leq 2 ; 1 \leq j \leq b ; 1 \leq k \leq n_{ij};$$

$$n_{1j} = 1 , n_{2j} = 1 \text{ or } 2 \dots \text{ or } n .$$

Remark * 1 leads us to consider the expectation of X_{ijk} as following

$$E(X_{ijk}) = \mu + \alpha_i + \beta_j + \theta_{ij} \quad \text{where the possible presence of interaction is taken into account with } \theta_{ij} .$$

Remark * 2 before further definition of the model, we can notice that we are in the frame of a classical design with proportional numbers of repetitions.

Table of numbers of repetitions by combination "block x density".

Block Density	1		j		b	Total
d_1	1		1		1	b
d_2	n		n		n	bn
Total	1 + n		1 + n		1 + n	b(1 + n)

Here are the conditions applying to the parameters θ , α and β .

$$\forall j \quad \sum_i n_{ij} \theta_{ij} = 0 \implies \forall j , \quad \theta_{1j} + n\theta_{2j} = 0$$

$$\forall j \quad \sum_i n_{ij} \theta_{ij} = 0 \implies \begin{cases} \text{for } i = 1 & \sum_j \theta_{1j} = 0 \\ \text{for } i = 2 & \sum_j \theta_{2j} = 0 \end{cases}$$

$$\text{while} \quad \sum_i \left(\sum_j n_{ij} \right) \alpha_i = 0 \implies \alpha_1 + n \alpha_2 = 0$$

$$\sum_j \left(\sum_i n_{ij} \right) \beta_j = 0 \implies \sum_{j=1}^b \beta_j = 0$$

Table of analysis of variance and formulas

Sources of variations	sums of squares	Degrees of freedom	Mean sums of squares	Computed F
Block	SSB	b-1	MSB	F _B
Density	SSA	1	MSA	F _A
Interaction	SSAB	b-1	MSAB	F _{AB}
Residual	SSE	b(n-1)	MSE	
Total	TSS	b(n+1)-1		

$$SSE = \sum_{ijk} (x_{ijk} - x_{i\cdot\cdot})^2 = \sum_{jk} (x_{2jk}^2) - \frac{\left(\sum_{jk} x_{2jk}\right)^2}{n},$$

$$MSE = \frac{SSE}{b(n-1)},$$

$$SSB = \sum_{j=1}^b (n-1)(x_{\cdot j\cdot} - x_{\cdot\cdot\cdot})^2 = \sum_{j=1}^b \frac{\left(\sum_{ik} x_{ijk}\right)^2}{n+1} - \frac{\left(\sum_{ijk} x_{ijk}\right)^2}{b(n+1)},$$

$$MSB = \frac{SSB}{b-1}, \quad F_B = \frac{MSB}{MSE},$$

$$SSA = \sum_{i=1}^2 \left(\sum_j n_{ij}\right) (x_{i\cdot\cdot} - x_{\cdot\cdot\cdot})^2 = \frac{\left(\sum_{jk} x_{1jk}\right)^2}{b} + \frac{\left(\sum_{jk} x_{2jk}\right)^2}{nb} - \frac{\left(\sum_{ijk} x_{ijk}\right)^2}{b(n+1)}$$

$$MSA = SSA, \quad F_A = \frac{MSA}{MSE},$$

$$SSAB = TSS - SSA - SSB - SSE \quad \text{and} \quad MSAB = \frac{SSAB}{b-1},$$

$$\text{where } TSS = \sum_{ijk} x_{ijk}^2 - \frac{\left(\sum_{ijk} x_{ijk}\right)^2}{b(n+1)}, \quad F_{AB} = \frac{MSAB}{MSE}.$$

Remark * 3 These formulas are only consequences of the generalized linear model.

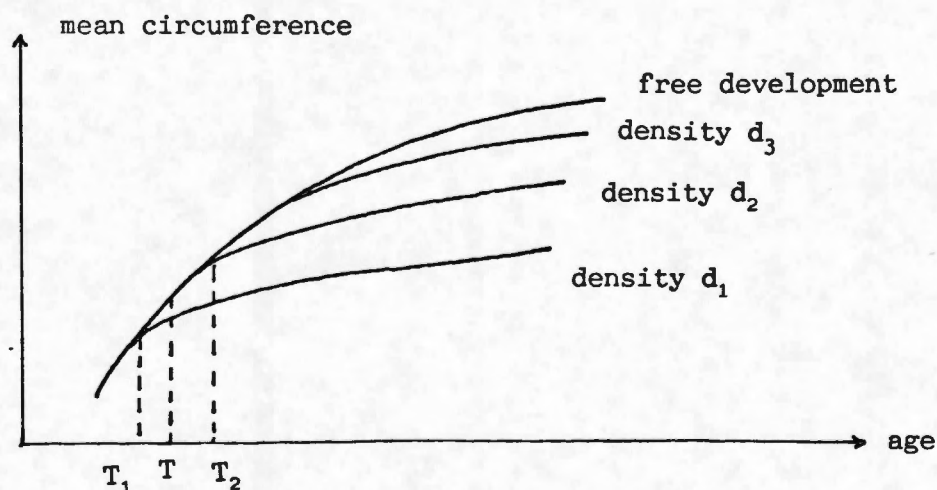
Numerical example : CCT.plots of the Bayottes (SENEGAL)

Table of data (the studied variable is the circumference)

Block	Density d_1	Density d_2						
1	31.30	34.15	35.38	30.37	32.06	31.68	27.63	31.95
2	31.96	34.21	32.03	32.90	33.07	34.63	34.13	32.80
3	28.79	33.60	32.80	30.22	29.51	31.41	29.17	32.04
4	31.08	30.22	32.42	32.58	30.97	32.57	33.52	35.21

Table of the analysis of variance

Sources of variation	Sums of Squares	Degrees of freedom	Mean sums of squares	F
Block	21.70592	3	7.2353	2.28
Density	7.62205	1	7.62205	2.40
Interaction	1.54	3	0.51	0.16
Residual	76.05	24	3.17	
Total	106.91797	31		



In order to understand the interests of the CCT.plots, we shall remind the aims of this trial. Its purpose is to compare the densities d_1 and d_2 and to detect the moment when competition arises in d_1 but is still away from d_2 in order to let some plots of the d_2 density to d_3 density (using clearing) for them to follow up their "free development". The moment of intervening T , should be situated between T_1 and T_2 , but as near as possible of T_2 . On the other hand, intervening after T_2 has more serious consequences.

Deciding intervention (clearing) is a consequence of the tests associated to the analysis of variance ; therefore this remark has an influence on the choice of the value of the significant level of the test α .

Test of interaction :

One has to test $H_0 : \forall ij, \theta_{ij} = 0$

against $H_1 : \exists ij, \theta_{ij} \neq 0$

Here $F_{AB} = 0.16$, this leads to accept the H_0 hypothesis. Anyway, it does not mean that there is no difference between d_1 and d_2 . On the other hand, if F_{AB} has been so that it would had led to rejection of H_0 , one would have decided that competition was get present in some blocks, and therefore a further study of θ_{ij} would have been necessary.

Test of density effect :

One has to test $H_0 : \forall i, \alpha_i = 0$

against $H_1 : \alpha_1 < \alpha_2 \quad (d_1 > d_2)$

The structure of H_1' is important, and we have to point out that H_1' is not expressed as $\exists i \alpha_i \neq 0$. This leads not to test the value of F_A but to consider the statistic :

$$T = \frac{X_{1..} - X_{2..}}{\sqrt{V \left(\frac{1}{b} + \frac{1}{nb} \right)}}$$

which under H_0 follows a Student distribution with 24 degrees of freedom in the actual case.

Notations : V is a random variable which is equal to $\hat{\gamma}$ estimate of γ .

$X_{1..}$ and $X_{2..}$ are random variables which are equal to the means calculated on the plots where the densities are respectively d_1 and d_2 .

$$\text{Here } x_{1..} - x_{2..} = \frac{123.13}{4} - \frac{903.28}{28} = 1.48$$

$$\hat{\gamma} = 3.17$$

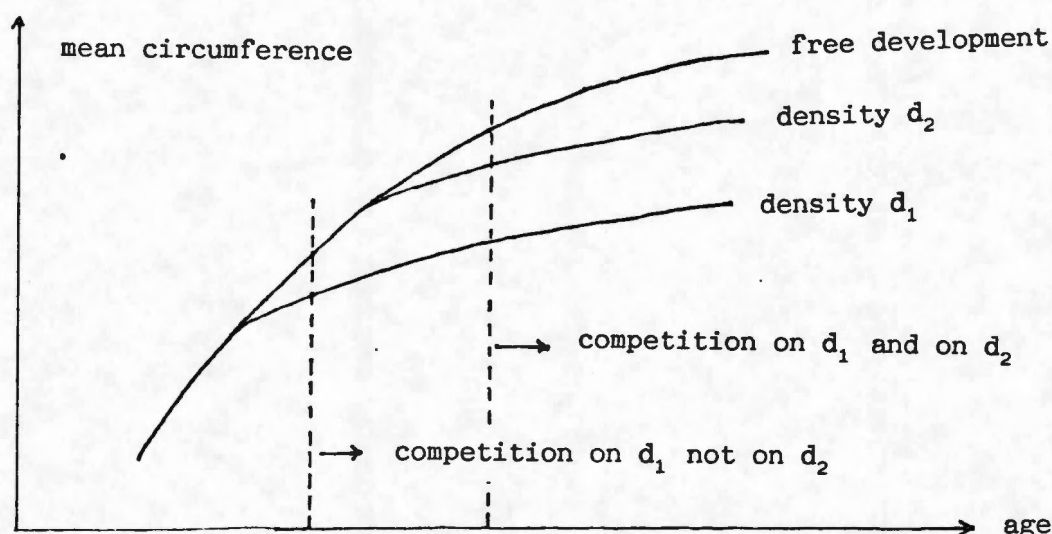
$$\text{then } t = \frac{x_{1..} - x_{2..}}{\sqrt{\hat{\gamma} \left(\frac{1}{b} + \frac{1}{nb} \right)}} = \sqrt{F_A} = \sqrt{2.40} = 1.63$$

H_0' rejected when $t > t_{\alpha, 24}$; considering remark * 4 , α can be chosen high. ($\alpha = 0.10$ rather than $\alpha = 0.05$ which is usually used).

When $\alpha = 0.10$, $t_{\alpha, 24} = 1.10$, so here H_0 is rejected.

Conclusion : It is decided to make a clearing in the d_2 plots.

Remark * 5 rejecting H_0' means that competition has started on plots of d_1 density, but it does not mean that competition has not begun on plots of d_2 density (see figure below).



III - The analysis of covariance :

The analysis of covariance is often proposed as a method of analysing the CCT plots instead of the analysis of variance discussed in II. One should anyway remember that the models used in this case are seldom justified and that the analysis of covariance is very far from having the same robustness properties that the analysis of variance.

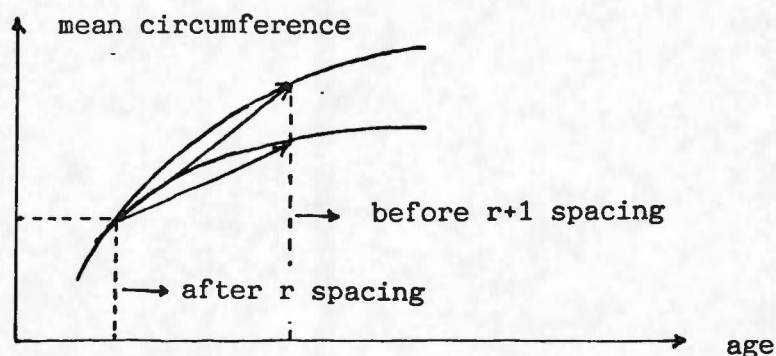
Model : The analysis of covariance differs from the analysis of variance, because it takes in account the measures $X_{ijk}^{(1)}$ after the r^{th} clearing to analyse the measures $X_{ijk}^{(2)}$ before the $(r+1)^{th}$ clearing.

For a start, we can write :

$$X_{ijk}^{(2)} = \mu + \alpha_i + \beta_j + \theta_{ij} + \lambda X_{ijk}^{(1)} + E_{ijk} .$$

The calculus are easy to lead (see appendix B for numerical example of data processing), but this model is not the best one can think about.

The figure hereafter shows that the effect of the initial state ($X^{(1)}$) depends on the density of the population.



So an appropriate model could be written as following :

$$E(X_{ijk}^{(2)}) = \mu + \alpha_i + \beta_j + \theta_{ij} + \lambda_i X_{ijk}^{(1)} .$$

One following model could be considered too :

$$E(X_{ijk}^{(2)}) = \mu + \alpha_i + \beta_j + \theta_{ij} + \lambda_{ij} X_{ijk}^{(1)} .$$

but it is very difficult to analyse.

Even in case of the preceeding model, it is not so easy to perform the analysis.

Then a good alternative may be to consider the simple analysis of variance led on the variable "increase of circumference between r and $r+1$ spacings". (See appendix B for numerical example and data processing).

The model is :

$$Z_{ijk} = \mu + \alpha_i + \beta_j + \theta_{ij} + E_{ijk} ,$$

where Z represents the increase of circumference, that is to say :

$$Z_{ijk} = X_{ijk}^{(2)} - X_{ijk}^{(1)} .$$

APPENDIX : Numerical example concerning the CCT-plots

This data set concerns a study of the growing of Teck in the Bayottes forest in Senegal

Plot n°	Density	Block n°	Mean circumference in 1981 = $X^{(1)}$	Mean circumference in 1980 = $X^{(2)}$
1	2	1	56.74	55.53
2	2	1	55.09	55.42
3	2	1	53.78	52.41
7	2	2	50.15	48.99
8	2	2	52.69	51.41
10	2	2	48.33	47.33
11	2	2	50.66	49.63
14	2	1	52.26	50.79
16	2	4	51.79	50.06
17	2	1	50.31	48.79
19	2	3	49.10	48.01
20	2	3	51.68	50.71
21	2	2	52.51	51.50
22	2	3	49.46	48.72
24	2	3	52.09	50.93
25	2	4	52.51	50.99
26	2	4	55.85	54.26
27	2	4	54.50	53.25
29	2	4	50.74	49.37
31	2	3	55.71	54.38
5	1	2	53.98	51.68
13	1	1	51.06	50.15
28	1	4	51.97	50.57
32	1	3	50.19	49.47

I Analysis of covariance of $X^{(2)}$ with $X^{(1)}$ covariate (without interaction).

Analysis of covariance table

The model is $X_{ijk}^{(2)} = \mu + \alpha_i + \beta_j + \lambda X_{ijk}^{(1)} + E_{ijk}$

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
trait	.07315	1	.07315	.33	.5753
bloc	.92901	3	.30967	1.38	.2814
mes80	94.29339	1	94.29339	419.73	.0000
ERROR	4.04372	18	.22465		

II How to get this analysis of covariance ?

Here we give a method to get this analysis of covariance without analysis of covariance program :

1. Models with treatment effect and block effect.

a) First, we compute the analysis of variance of $X^{(2)}$:

The model is : $X_{ijk}^{(2)} = \mu^o + \alpha_i^o + \beta_j^o + E_{ijk}^o$

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
trait	.82502	1	.82502	.16	.6942
bloc	17.04529	3	5.68176	1.10	.3743
ERROR	98.33711	19	5.17564		

b) Then we compute the same analysis with dependent variable $X^{(1)}$

The model is : $X_{ijk}^{(1)} = \mu^* + \alpha_i^* + \beta_j^* + E_{ijk}^*$

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
trait	1.43665	1	1.43665	.28	.6022
bloc	16.76452	3	5.58817	1.09	.3763
ERROR	97.16464	19	5.11393		

c) By computing the regression of E_{ijk}^o against E_{ijk}^* we get :

$$E_{ijk}^o = E_{ijk}^* + E_{ijk}$$

The table of analysis of variance associated to this regression is :

	SUM OF SQUARES	DF	MEAN SQUARE
REGRESSION	94.3004	1	94.3004
RESIDUAL	4.0522	18	

Remark : The degree of freedom of residuals is 18 and not 22 ($23 - 4 - 2 + 1 = 18$), because we have some relations between residuals. Another point, the regression passes through the origin.

Note that the sums of squares associated to the regression and the residual of this table are equal to those associated to the covariate and the residual of the table of analysis of covariance.

2. Models with block effect without considering treatment effect.

In order to get the others sums of squares of the analysis of covariance, we have to study the following models the same way :

a) $X_{ijk}^{(2)} = \mu^{ot-} + \beta_j^{ot-} + E_{ijk}^{ot-}$

The notation with index "t-" indicates that we do not consider any treatment effect in these models.

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
bloc	17.04529	3	5.68176	1.15	.3548
ERROR	99.16213	20	4.95811		

$$b) X_{ijk}^{(1)} = \mu^{*t-} + \beta_j^{*t-} + E_{ijk}^{*t-}$$

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
block	16.76452	3	5.58817	1.13	.3595
ERROR	98.60128	20	4.93006		

$$c) E_{ijk}^{ot-} = \lambda^{t-} E_{ijk}^{*t-} + E_{ijk}^{t-}$$

	SUM OF SQUARES	DF	MEAN SQUARE
REGRESSION	95.0450	1	95.0450
RESIDUAL	4.1173	19	

The sum of squares associated to the treatment in the complete model equals the residual sum of squares of the model :

$$E_{ijk}^{ot-} = \lambda^{t-} E_{ijk}^{*t-} + E_{ijk}^{t-}$$

minus the residual sum of squares of the

$$E_{ijk}^{o-} = \lambda E_{ijk}^{*} + E_{ijk}$$

model.

We can verify that : $0.07314793 = 4.11684754 - 4.04369961$

3. Models with treatment effect without considering block effect.

By writing the same kind of tables considering the model without blocks we get :

$$a) X_{ijk}^{(2)} = \mu^{ob-} + \alpha_i^{ob-} + E_{ijk}^{ob-}$$

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
trait	.82502	1	.82502	.16	.6955
ERROR	115.38240	22	5.24465		

$$b) X_{ijk}^{(1)} = \mu^{*b-} + \alpha_i^{*b-} + E_{ijk}^{*b-}$$

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
trait	1.43665	1	1.43665	.28	.6037
ERROR	113.92915	22	5.17860		

$$c) E_{ijk}^{ob-} = \lambda^{b-} E_{ijk}^{*b-} + E_{ijk}^{b-}$$

	SUM OF SQUARES	DF	MEAN SQUARE
REGRESSION	110.4094	1	110.4094
RESIDUAL	4.9730	21	

Again, we can verify that :

$$0.92902115 = 4.97272076 - 4.04369961$$

We have produced these tables just to point out the method to produce the table of the analysis of covariance in a simple case.

III Analysis of covariance of $X^{(2)}$ with $X^{(1)}$ covariable (with interaction).

Applying the same method to a more complicated model :

$$X_{ijk}^{(2)} = \mu + \alpha_i + \beta_j + \theta_{ij} + \lambda X_{ijk}^{(1)} + E_{ijk}$$

we can write the following table of analysis of covariance.

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.	REGRESSION COEFFICIENTS
trait	.05264	1	.05264	.30	.5903	
bloc	1.52951	3	.50984	2.93	.0676	
tb	1.43557	3	.47852	2.75	.0791	
mes80	80.25507	1	80.25507	461.56	.0000	.95077
ERROR	2.60815	15	.17388			

And now, here are the results concerning the analysis of variance applied to the increase of circumference. The model is :

$$Z_{ijk} = X_{ijk}^{(2)} - X_{ijk}^{(1)} = \mu + \alpha_i + \beta_j + \theta_{ij} + E_{ijk}.$$

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
trait	.08427	1	.08427	.48	.4994
bloc	1.47614	3	.49205	2.79	.0743
tb	1.24188	3	.41396	2.35	.1114
ERROR	2.82337	16	.17646		